EXHIBIT 20

To: Amy Buono, Gerry Dillon, PhD

From: Steven Haist, MD, MS, FACP

Re: Comparison of Step 2 forms C01607and C02385

I have been asked to render a judgement regarding whether the above two Step 2 forms are comparable or not. The analyses were run by Technology, Training and Support staff (Deniz Bucak) and others involved in acquiring the information included Peg Johnson, Francine Rosenthal and Kathleen Short).

The Step 2 examination in 2007 was 368 items (46/hour) of which 288 items were live; in 2011, the Step 2 examination was 352 items (44/hour or 42/hour, if the hour block contained a scientific abstract or pharmaceutical advertisement, the block contained 2 fewer items) of which 280 are live.

The content below, only addresses scored items, and does not include pretest items. The Step 2 examination currently has 177 different exam content specifications, the examination in 2007, also had 177 (Patient Safety was added since 2007, and Heated Related Illness was combined with Pituitary, hypothalamic disorders, both of which were under Endocrinology. Thus instead of one item being from both Heat related illness **and** one from Pituitary and hypothalamic disorders, 1 item is from Heat related illness **or** one from Pituitary and hypothalamic disorders. Of the 177 categories in 2007, specifications changed in 18; because of the reduction in scored items on the examination, 288 to 280, the 8 items had to be removed from content categories, and this reduction was proportional as is our practice whenever there is such a change in specifications. Thus the change in the number of content categories in which there was a change was 10 (5.6%) for reasons unrelated to timing. Of the 18 content categories where there was a change, 14 of content categories decreased by 1 (6 of these went from 2 in 2007 to 1 in 2011; 2 went from 3 to 2; 3 went from 5 to 4; 1 went from 6 to 7 and 1 went from 7 to 8; and there was one content category, Heat Related Illness that was combined with another content area, Pituitary, hypothalamic disorders).

For 4 of the content categories, the number of items increased, 3 of them by 1 (Patient Safety had 2 subcategories and since they were new, the number of items in each went from 0 to 1; and interpretation of the literature increased from 3 to 4 items). Understanding statistical concepts of measurement in medical practice increased from 2 to 5. The Step 2 Committee approved this change in response to the Comprehensive Review of USMLE (CRU) recommendation of "introduce a testing format designed to assess an examinee's ability to obtain, interpret, and apply scientific and clinical information needed to solve a clinical problem, i.e., to engage in evidence-based decision making."

By content specifications, the 2007 to 2011 examination were essentially the same (93.8% of the items in the 2007 examination were from the same specified content category as items in the 2011 examination) and there was only 1 new category (2 items in Patient safety, 2 subcategories, 1 item in each subcategory). The 93.8% should actually be higher since the number of items on the examination decreased by 8; the adjusted percentage of items that were from the same content category is 96.4%.

There are 4 task codes; preventive, mechanisms of disease, diagnosis and management. The task code relates to the objective of the question (e.g. What is the most likely diagnosis?) The specifications in 2007 and 2011 for the categories were as follows: preventive, 32 in 2007 and 31 in 2011 (a change of 3.1%); mechanisms of disease, 65 in 2007 and 63 in 2011 (a change of 3.1%); diagnosis, 109 in 2007 and 104 in 2011 (a change of 4.6%); and management, 63 in 2007 and 59 in 2011 (6.3%). Again, these percentages need to be adjusted because of the reduction in the number of items on the 2 forms (2007 and 2011). The adjusted percentages would be 0%, 1.6%, 1.8% and 1.6%, respectively. There is no significant difference by task code between the 2007 and 20011 forms.

The types of MCQs did change as one might expect as an assessment organization evaluates new methodology and new technology but this change is of little importance. The single best answer with an option set after the text (A-type) was 20 more in 2011 (248 in 2007 and 264 in 2011); there were 40 R-types in 2007 (two or more items about a common theme with an extended option set presented first and then a vignette describing a particular patient presented after the option set) vs. 12 in 2011. In 2011 there were 4 F-types and 0 in 2007. F-types are set of single best answer items that unfold over time. A type, R-type and the F-type are all single best answer thus little or inconsequential significance in the test forms.

The number of items with pictures increased as well. In 2007, 14 of 288 (4.9%) had associated pictures (e.g. Computerized Tomography [CT] images, x-rays, and electrocardiograms). The 2011 examination had 25 pictures (8.9%). Again, one would expect the number of images, photographs, electrocardiograms to increase on a medicine certifying examination keep pace with technology and advances in assessment. The targeted range in items for the 2007 examination was a minimum of 5 to a maximum of 16 and for the 2011 examination the targeted range was 21-36. There were 3 items (1.0% of the examination) that required the identification of heart sounds or murmurs in 2011. Within the item was an associated video of an avatar and a stethoscope that could be moved to different locations on the chest. The examinee would listen through ear phones to the particular heart sound. These items and the required technology have been introduced to the USMLE since the 2007 examination.

Overall, the 2007 Step 2 form STP2C01607 and the 2011 Step 2 for, STP2C02385 are very similar with differences which I would consider minimal. The changes noted above (more pictures and F-types and the inclusion of heart sounds) are changes consistent with the evolution of an examination over 4 years. The content is essentially the same except for the addition of 2 Patient Safety items and an increase in the number of Biostatics and epidemiology items.

If you have any other questions please contact me.